AMD 34

CONFIRMATION

CLAIMS

SUB A1

1. A set of libraries of genes which code for proteins which are capable of specific binding interactions with a specific binding partner by amino acid residues at at least two specified positions including a first specified position and at least one other specified position, which set of libraries consists of:

a) 6 to 20 libraries in which each library has a triplet that codes for at least one but less than 20 amino acids at said first specified position, and is randomised at the or each triplet coding for the said at least one other specified position, the arrangement being such that interactions of the proteins coded for by the said 6 to 20 libraries with a specific binding partner identifies a triplet that codes for an amino acid at the said first specified position that takes part in the specific binding interaction, and

b) 6 to 20 libraries in each of which libraries said first specified position is randomised and a different one of said at least one other specified positions has a triplet that codes for at least one but less than 20 amino acids.

2. The set of libraries of genes as claimed in claim 1, which set of libraries consists of:

a) 12 libraries in which each library has a triplet that codes for one or several but less than 20 amino acids at the said first determined position, the triplets being as shown in Table 1 or Table 2, and

b) 12 libraries of corresponding design for each of the said one or more other determined positions.

3. The set of libraries of genes as claimed in claim 1 or claim 2, wherein the genes code for zinc fingers.

CONFIRMATION

4. The set of libraries of genes as claimed in claim 3, which set consists of 36 libraries in three groups of 12 libraries which code for amino acids at the -1 and +3 and +6 positions respectively.

5. The set of libraries of genes as claimed in claim 3 or claim 4, wherein each gene codes for a protein comprising 3 zinc fingers.

6. The set of libraries of genes as claimed in claim 5, wherein each gene codes for a protein having the sequence (SEQ ID NO: 2)

T G E K P Y K C P E C G K S F S X K S X L V X H Q R T H

T G E K P Y K C P E C G K S F S X K S X L V X H Q R T H

T G E K P Y K C P E C G K S F S X K S X L V X H Q R T H.

where X is any amino acid

7. A set of libraries of proteins, which proteins are capable of specific binding interactions with a specified binding partner by amino acid residues at at least one specified position including a first specified position and at least one other specified position, which set of libraries consists of:

a) 6 to 20 libraries in which each library has at least one but less than 20 amino acid residues at the said first specified position and is randomised at the said at least one other determined position, the arrangement being such that interaction of the 6 to 20 libraries with a specific binding partner identifies an amino acid residue at the said first specified position that takes part in the specific binding interaction, and

b) 6 to 20 libraries in each of which libraries said first specified position is randomised and a different amino acid is present at at least one other specified position.

CONFIRMATION

8. The set of libraries of proteins as claimed in claim 7, which set of libraries consists of

a) 20 libraries in which each library has one specified amino acid residue at the said first determined position and is randomised at the said one or more other determined positions, and

b) 20 libraries of corresponding design for each of the said one or more other determined positions.

9. The set of libraries of proteins as claimed in claim 7 or claim 8, wherein the proteins are zinc fingers.

10. The set of libraries of proteins as claimed in claim 7, which set consists of 60 libraries in three groups of 20 libraries with specified amino acids at the -1 and +3 and +6 positions respectively.

11. The set of libraries of proteins as claimed in claim 9 or claim 10, wherein each protein comprises three zinc fingers.

12. The set of libraries of proteins as claimed in claim 11, wherein each protein has the sequence (SEQ ID NO: 2)

T G E K P Y K C P E C G K S F S X K S X L V X H Q R T H

T G E K P Y K C P E C G K S F S X K S X L V X H Q R T H

T G E K P Y K C P E C G K S F S X K S X L V X H Q R T H.

where X is any amino acid

13. A set of libraries of genes which code for the set of libraries of proteins defined in any one of claims 7 to 12.

CONFIRMATION

14. A method of identifying a protein which interacts with a specific binding partner, which method comprises providing a set of libraries of proteins as defined in any one of claims 7 to 12, incubating the specific binding partner with each library of the set, observing specific binding interactions with certain libraries of the set, and using the observations to identify a protein which interacts with the specific binding partner.

15. The method as claimed in claim 14, wherein the specific binding partner is a polynucleotide.

16. The method as claimed in claim 14, wherein the specific binding interactions are observed by radiometric or luminescent assay.

17. The method as claimed in claim 14, wherein the specific binding interactions are observed by imaging means.

18. The method as claimed in claim 14, wherein the specific binding interactions are observed by scintillation proximity assay.

19. The method as claimed in claim 18, wherein the sets of libraries of proteins are immobilised on scintillation proximity assay surfaces and the specific binding partner is radiolabelled.

20. The method of claim 18 or claim 19, wherein after incubation the scintillation proximity assay surfaces are washed to distinguish stronger specific binding interactions from weaker ones.

CONFIRMATION

21. The method as claimed in claim 14, wherein the specific binding interactions are observed by colorimetric means.

22. The method as claimed in claim 21, wherein the specific binding partner is biotinylated and the specific binding interaction is detected using a signal generating streptavidin conjugate.

23. The method as claimed in claim 21 or claim 22 wherein after incubation the binding interactions are washed to distinguish stronger specific binding interactions from weaker ones.

SUB A2

24. A protein having the sequence (SEQ ID NO: 1)

T G E K P Y K C P E C G K S F S . K S ✣ L V $ H Q R T H

T G E K P Y K C P E C G K S F S . K S ✣ L V $ H Q R T H

T G E K P Y K C P E C G K S F S . K S ✣ L V $ H Q R T H.

25. A gene which codes for the protein of claim 24.

26. A method of constructing randomised gene libraries in which the number of genes is the same as the number of encoded proteins and which contain no termination codons at the predetermined positions of randomisation, the method comprising the steps of:

a) providing a template oligonucleotide which is fully randomised at predetermined codon positions;

b) for each predetermined codon position providing a pool of selection oligonucleotides, wherein each member of said pool contains a different codon selected from the group consisting of

AAA, AAC, ACC, AGC, ATG, ATT, CAG, CAT, CCG, CGC, CTG, GAA, GAT, GCG, GGC, GTG, TAT, TGG, TGC, TTT.

at the predetermined codon position;

c) selecting one or more selection oligonucleotides from each pool in order to encode the required gene or library;

d) allowing the ligated selected oligonucleotides from each pool to hybridise with the template oligonucleotide;

e) forming one or more constructs by ligating the hybridised selection oligonucleotides together;

f) removing a region from a gene of interest corresponding to the hybridised product from step e);

g) forming a gene library or genes by ligating the products from step e) into the said gene of interest wherein the said gene of interest is contained within a suitable expression vector.

27. A method of producing proteins encoded by the randomised gene libraries of claim 26 comprising the steps of:

a) transforming a suitable host cell with the gene or gene library of claim 27 construct;

b) expressing the genes to form proteins;

c) purifying the proteins.